

In the Title:

Please amend the title of the application as follows:

Group-By Result Size Result Estimation

In the Specification:

Please amend the paragraph beginning on page 4 line 21 as follows:

~~FIG. 3 FIGS. 3A and 3B are~~ is a flow ~~chart~~ charts illustrating a method of calculating a selectivity factor according to the preferred embodiment of this invention, and is suggested for printing on the first page of the issued patent.

Please amend the paragraph beginning on page 8 line 15 as follows:

However, there are limitations associated with the illustrated prior art method of calculating selectivity. For example, the prior art does not take advantage of the relationship of columns in the same table. Nor does the prior art consider equivalent columns that are not explicitly specified in the Group-By operation. The presence of these limitations produces an inaccurate estimation of selectivity as it is not a true reflection of the relationship of the columns in the operation. ~~Fig. 3 is~~ Figs. 3A and 3B are a flow chart (100) illustrating a novel method of calculating selectivity according to the present invention which is used to calculate a more accurate estimate of the result size of the Group-By operation in situations where there are multiple Group-By columns, and more so when the operation involves multiple columns from two or more tables. A Group-By operation often involves the use of multiple tables, wherein different tables are different sizes, *i.e.* different numbers of rows. To compensate for the inclusion of multiple tables in the operation, a factor is applied to each selectivity for each column. The first subroutine in the process of calculating selectivity is to find the table in the operation with the largest quantity of rows, T_{max} , *i.e.* the largest size table. Initially T_{max} is set at a value of one (102). Thereafter, for each table ‘T’ in the operation (104), the table size based upon the number of rows, T_i , for that table is obtained (106). The table size T_i is compared to the T_{max} value, and the larger table size value between the two values is set as T_{max} (108). Thereafter, a test is conducted to determine if the table considered at step (108) is the last table in the operation (110). If the response to the test at step (110) is negative, ‘T’ is incremented and the subroutine returns to step (106) to obtain the table size for the next table in the operation. However, if the response to the test at step (110) is positive, the subroutine is complete and the final T_{max} value is the size of the largest table in the operation based upon the criteria of the number of rows in the tables.

Please amend the paragraph beginning on page 11 line 1 as follows:

The process shown herein for calculating the selectivity for the Group-By operation takes into consideration a plurality of factors that are not present in the prior art, wherein these factors provide an accurate estimation of the Group-By selectivity. Once the calculation of the Group-By selectivity has been completed, *i.e.* a positive response to the test at step (128), the product of the Group-By selectivity and the input size of the operation, as shown at step 22 of Fig. 1, will result in an accurate estimate of the result size of the Group-By operation. Accordingly, the process outlined in Fig. 3 Figs. 3A and 3B demonstrate an improved method of calculating the selectivity in a Group-By operation that takes into consideration the maximum size of the tables in the operation, the relative uniqueness of a column based upon the size of the table in which the column resides, the equivalency of columns based on equal join predicates, and the probability of union of two or more columns from two or more tables.

Please amend the paragraph beginning on page 11 line 13 as follows:

The invention as shown in Figs. 1, 3A, and 3B and ~~3~~ is illustrated as a method. However, it may also be a module associated with a database or an article of manufacture embodied within the instructions of a computer implemented system. For example, the module may be a Group-By result size estimator used in a database system, or it may be embedded within the database system.